

Predicción del comportamiento diario de la acción de SURAMINV.

Un modelo de redes neuronales*

Prediction of the daily share price fluctuations of SURAMINV

A neural network model

Jaime Enrique Arrieta Bechara**

Juan Camilo Torres Cruz***

Hermilson Velásquez Ceballos, Ph.D.****

Recepción: Agosto 13 de 2009

Aceptación: Diciembre 02 de 2009

Resumen

La investigación muestra que es posible realizar, por medio de modelos de redes neuronales artificiales, buenas predicciones sobre el comportamiento diario de la acción de SURAMINV. Tales resultados contrarian la hipótesis de la teoría de eficiencia débil de mercado. A partir de dichas predicciones y con el uso de sistemas de negociación, se evalúa la posibilidad de obtener rendimientos extraordinarios sobre la estrategia *Buy & Hold* teniendo en cuenta costos de transacción y oportunidad.

Palabras clave: Red neuronal artificial (RNA), predicción, Suraminv, sistemas de negociación.

* Artículo derivado del trabajo de investigación desarrollado por los autores en la línea "La política coyuntural de la empresa" en el grupo de Estudios en Economía y Empresa, Universidad EAFIT.

** Magíster en Finanzas, Economista, Universidad EAFIT. Jefe Mesa Derivados Valores Bancolombia. jarriet1@eafit.edu.co

*** Magíster en Finanzas, Universidad EAFIT. Especialista en Ingeniería Financiera, Universidad Nacional de Colombia. Economista del Desarrollo, Universidad Pontificia Bolivariana. jtorres7@eafit.edu.co

**** Dr. Ciencias Matemáticas, Universidad Politécnica de Valencia. Magíster en Matemáticas Aplicadas, Universidad EAFIT. Profesor de tiempo completo, Universidad EAFIT. evelas@eafit.edu.co

Abstract

As opposed to the weak form of efficient-market hypothesis, the current study shows that it is possible to do good predictions about the daily share price fluctuations of Suraminv, using artificial neural network models. Furthermore, the forecasts obtained are used to analyze the possibility of gaining extraordinary returns with regard to the *Buy & Hold* strategy, through negotiation systems with basic rules.

Key words: Artificial Neural Network (ANN), prediction, Suraminv, negotiation systems.

Introducción

La teoría de Eficiencia de Mercado constituye uno de los pilares del desarrollo moderno de las finanzas. Según esa teoría, y como lo explica Fama (1970), los precios de las acciones son el reflejo de toda la información disponible. Por tal razón, ningún agente podría obtener retornos extraordinarios en los mercados de manera consistente.

Roberts (1967) distingue tres tipos diferentes de eficiencia en los mercados: débil, semifuerte y fuerte. Se distinguen por la clase de información con la que cuentan los agentes. La hipótesis de la eficiencia débil es que las series históricas del mercado, que incluyen precios, volúmenes y demás datos referentes a las transacciones, no poseen ningún tipo de información que pueda ser aprovechable por los agentes para obtener rendimientos extraordinarios de forma consistente. Según este planteamiento, no es posible predecir precios ni rendimientos futuros de los activos.

Se podría decir que el estudio de la predictibilidad de los activos financieros data de cuando empezaron los mercados donde aquellos se negocian. No obstante, hay en curso un proceso de perfeccionamiento y sofisticación de las herramientas empleadas para dichas predicciones, lo cual permite hoy la utilización de un amplio abanico de posibilidades para tal

fin. La literatura existente sobre predictibilidad en los mercados financieros y evaluación de estrategias de negociación es muy amplia. A continuación se mencionan los estudios más destacados a nivel internacional.

Hua y Yann (1996) han desarrollado un modelo de redes neuronales recurrentes para hacer predicciones a mediano plazo en el mercado accionario de Taiwán. Usaron entradas provenientes del análisis ARIMA de la serie de estudio. Sus resultados permitieron observar que, si se entrena la red con datos semanales con cuatro años de historia, se pueden lograr buenas predicciones del mercado para un horizonte de seis meses.

Fernández, González y Sosvilla (2000) aplican una regla de negociación para el Índice General de la Bolsa de Madrid con base en redes neuronales y análisis técnico. Estos autores concluyen que dicha regla obtiene mejores rendimientos de forma consistente frente a una estrategia *Buy & Hold* mientras el mercado se encuentra estable o a la baja; cuando la tendencia del mercado se encuentra en alza, la mejor estrategia es la pasiva. Por su parte, Chen, Leung y Daouk (2003) buscaron modelar y predecir la dirección de los retornos del Índice de la Bolsa de Acciones de Taiwán a través de estrategias de negociación basadas en redes neuronales probabilísticas; obtuvieron mejores rendimientos que con la estrategia *Buy & Hold*.

Parisi y Guerrero (2003) han hecho predicciones de los cambios de signo semanales de los índices bursátiles Bovespa (Brasil), CAC40 (Francia), Dow Jones Industrial (Estados Unidos), FTSE100 (Reino Unido), GDAX (Alemania), Hang Seng (Hong Kong), KLS (Malasia), Nikkei225 (Tokio), STI (Tailandia) y S&P500 (Estados Unidos), y han encontrado que esta técnica presenta mayor poder predictivo sobre las series estudiadas frente a modelos Arima.

Por su lado, Álvarez y Álvarez (2003) utilizan algoritmos genéticos y redes neuronales para predecir los precios del yen y la libra esterlina. Adicionalmente, hacen una fusión de datos a través de la cual combinan los resultados de sendos métodos con el fin de mirar la existencia de sinergias para mejorar las predicciones. Pese a lo sofisticado de las herramientas, no se observa, frente a trabajos anteriores, una mejora considerable en los resultados de predicción de tipos de cambio

Sallehuddin *et al.* (2007) proponen un modelo híbrido para la predicción de diferentes índices económicos y financieros asiáticos. El modelo sugerido es una combinación de redes neuronales tipo GRANN (*Grey Relational Artificial Neural Networks*) con un modelo lineal Arima. Para ellos, su modelo presenta una mejor predicción que los modelos con los cuales lo comparan.

La literatura sobre el tema con aplicación a mercados colombianos es también muy amplia. Se destacan trabajos como los de Pantoja (2000), donde se realiza una comparación de resultados de predicción a partir de modelos Arima y redes neuronales para once series de precios de mercados colombianos. Alcanzaron evidencia de mejor desempeño en las redes, pero detectaron que en la presencia de estacionalidad los modelos Arima son muy potentes. Gil y Pérez (2005), por su parte, realizan una comparación entre un modelo tipo Garch y uno de redes neuronales para la predicción de la acción de Acerías Paz del Río. Para ellos, al ser las redes

modelos más flexibles y adaptables, se pueden considerar como una buena alternativa para la predicción de series de tiempo, pero a su vez las consideran más como un complemento a las técnicas tradicionales que como un reemplazo.

Jalil y Misas (2007) hacen una comparación en la predicción del tipo de cambio a través del uso de modelos de redes neuronales y de tipo lineal. Reflejan mejores resultados de predicción con los primeros.

Villada *et al.* (2008) proponen un modelo de redes neuronales para el pronóstico del precio de la energía eléctrica colombiana. Las estructuras de redes a las que recurrieron tienen como entradas las series de precios diarios y el nivel medio de los embalses. Hacen una comparación con un modelo tipo Garch, el cual muestra un mejor desempeño de pronóstico dentro de la muestra, pero se observa que fuera de la muestra las redes tienen mejor predicción.

Como se puede observar, en la literatura sobre predicción de mercados existe una gran variedad de propuestas con base en diferentes técnicas. En cada caso se llega a concluir a favor o en contra de la teoría de eficiencia de mercado en forma débil. El presente estudio pretende ampliar la discusión sobre el tema y otorgar nuevos criterios que permitan contrastar la teoría financiera en mercados colombianos. Los trabajos realizados, donde se busca evaluar la eficiencia de los mercados, llegan a conclusiones favorables o contrarias a dicha teoría. Lo cierto es que en todo el mundo tiene mucha aceptación entre operadores de mercados financieros la implementación de estrategias que pretenden predecir comportamientos futuros a través de indicadores técnicos o modelos contruidos a partir de las series históricas de las transacciones.

La divergencia existente entre la teoría y la práctica justifica la realización de trabajos como el que se expone, en los se busque evaluar la

eficiencia de mercado en la negociación de activos específicos. Concretamente, es de gran relevancia evaluar la predictibilidad del comportamiento de la acción de Suramericana de Inversiones S.A. (Suraminv), ya que esta acción lideró las negociaciones en la Bolsa de Valores de Colombia durante el periodo de estudio, fue la de mayor participación en el Índice General de la Bolsa de Colombia (IGBC) y la que tuvo el mayor volumen de transacciones.

1. Redes Neuronales Artificiales (RNA)

Como las técnicas de aproximación lineal y polinomial, una red neuronal es un método que, por medio de un conjunto de algoritmos, busca establecer la relación entre un conjunto de variables de entrada, $\{x_i\}$, $i=1, \dots, n$, con un conjunto de una o más variables de salida, $\{y_j\}$, $j=1, \dots, k$. La diferencia de la red con los otros métodos de aproximación consiste en que la primera usa una o más capas ocultas, donde las variables de entrada son transformadas y comprimidas por una función especial. La red imita la forma en la cual el cerebro humano procesa datos sensoriales de entrada, recibidos como neuronas de entrada, hasta el reconocimiento como neuronas de salida (Mcnelis, 2005).

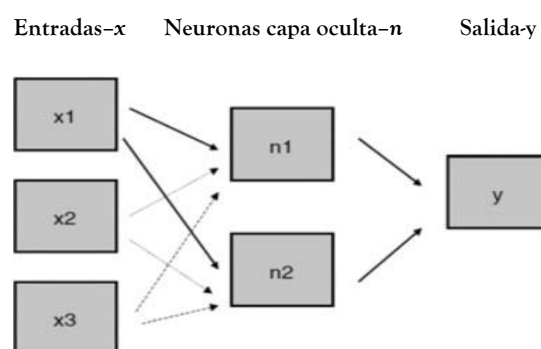
A la hora de realizar la modelación de variables financieras, la econometría tradicional ofrece modelos con especificaciones lineales y no lineales. Sin embargo, diversos autores resaltan ciertas bondades de las redes neuronales con respecto a estos modelos econométricos, por presentar un mejor ajuste y por tanto un menor error de predicción. En esta línea, se pueden mencionar trabajos como el de Franses y Dijk (2000). Estos autores resaltan las RNA frente a los modelos de series de tiempo tradicionales, por la posibilidad que ofrecen de modelar relaciones no lineales complejas. Por su parte, Shachmurove (2000) destaca como ventaja de las RNA el no requerir supuestos acerca de la

naturaleza de la distribución de los datos, y concluye que cumplen un mejor papel que los modelos de series de tiempo y econométricos cuando se trata de encontrar relaciones no lineales.

Aristizábal (2006) menciona que “la creciente popularidad de las RNA en el campo de la economía y las finanzas se debe a la necesidad de tratamientos econométricos distintos a los tradicionales, que sean capaces de capturar adecuadamente las trayectorias no lineales” de las relaciones entre variables.

Las redes *feedforward* son las más usadas en aplicaciones de economía y finanzas. Se caracterizan porque los datos se transmiten de las neuronas de capas anteriores a las de capas posteriores: las diferentes neuronas de la red procesan un conjunto de información de entrada para establecer la relación con la variable de salida. El modelo lineal es un caso especial de red *feedforward* en el que la única neurona de la capa oculta es una función de activación lineal que conecta la única capa de salida con un parámetro de peso. La gráfica 1 ilustra la arquitectura de una red neuronal *feedforward* con una capa oculta $\{x_i\}$, $i=1,2,3$, de dos neuronas, tres variables de entrada, y una salida y .

Gráfica 1
Red neuronal *feedforward*



Fuente: Elaboración propia

Las capas ocultas se pueden interpretar como el procesamiento que realizan los agentes de toda la información económica y financiera conocida, con el fin de formar sus expectativas hacia el futuro. Este procesamiento en el cerebro de los agentes, representado en las transformaciones realizadas en las capas ocultas de la red, es el que conlleva a reacciones y a tomar decisiones en los mercados donde se realizarán compras o ventas.

Las neuronas procesan los datos de entrada de dos formas: primero, haciendo combinaciones lineales de los datos de entrada y, luego, ejecutando esas combinaciones por medio de funciones de activación o transferencia. Las funciones más comunes en economía y finanzas son las no lineales compresoras, las cuales reportan salidas no proporcionales a las entradas y son, además, las que mejor representan el comportamiento económico de los agentes. Por su parte, para la capa de salida, la función de transferencia más usada es la lineal por ser continua en todos los reales. Arroja, como resultado, exactamente el mismo valor de la entrada neta que le llega y, por tanto, es equivalente a no aplicar función de transferencia.

Una de las funciones de activación compresora más usadas en este tipo de trabajos es la tangente hiperbólica, también conocida como la función *tansig* o *tanh*. Esta comprime las combinaciones lineales de las entradas en el rango comprendido entre -1 y +1 y está representada matemáticamente por:

$$n_{k,t} = \omega_{k,0} + \sum_{i=1}^{i^*} \omega_{k,i} x_{i,t} \quad (1)$$

$$N_{k,t} = T(n_{k,t}) \quad (2)$$

$$= \frac{e^{n_{k,t}} - e^{-n_{k,t}}}{e^{n_{k,t}} + e^{-n_{k,t}}} \quad (3)$$

$$y_t = \gamma_0 = \sum_{k=1}^{k^*} \gamma_k N_{k,t} \quad (4)$$

Donde $T(n_{k,t})$ es la función de activación *tansig* para la neurona de entrada $n_{k,t}$.

En este sistema hay i^* variables de entrada $\{x\}$ y k^* neuronas. Una combinación lineal de estas variables de entrada, observadas en el momento t , $\{x_{i,t}\}$, $i=1, \dots, i^*$, con el vector de coeficientes o conjuntos de pesos de entrada $\{\omega_{k,i}\}$, $i=1, \dots, i^*$, y con la constante $\omega_{k,0}$, forma la variable $n_{k,t}$. Esta variable es comprimida por la función *tansig* y se convierte en una neurona $N_{k,t}$ en el momento de la observación t . El conjunto de k^* neuronas en el momento t las combina en una forma lineal con el vector de coeficientes $\{\gamma_k\}$, $k=1, \dots, k^*$, y con un término constante γ_0 , para formar la predicción \hat{y}_t , en el momento t .

Por otra parte, para la capa de salida, la función de transferencia más usada es la lineal (*purelin*) por ser continua para todos los reales, pues arroja como resultado exactamente el mismo valor de la entrada neta que le llega; por tanto, esta operación es equivalente a no aplicar función de transferencia. La continuidad es una característica deseada para la función de esta capa en el sentido que la red tendrá libertad para arrojar cualquier valor de salida.

En los problemas de predicción económica o financiera se utiliza normalmente el algoritmo de entrenamiento *backpropagation*¹. Es un método de aprendizaje con un mecanismo de propagación-adaptación en dos fases: la primera consiste en entrar un determinado patrón y propagarlo por todas las capas de la red hasta llegar a una salida, comparada con la salida que se desea obtener; así se determina un término de error. Con este se inicia la segunda fase, ya que se convierte en una señal que se propaga hacia atrás, actualizando todos los pesos de la red. La explicación de su generalizada implementación en problemas económicos y financieros radica en las virtudes que posee: es capaz de resolver

¹ Autores como Wong (1995), y Yao, Li y Tan (1997) han documentado la superioridad de este tipo de red en el campo de economía y finanzas.

problemas que no son linealmente separables, reduce el tiempo de entrenamiento de la red por la utilización del procesamiento paralelo para encontrar la relación entre patrones dados y no necesita la especificación de un algoritmo por anticipado para llegar al algoritmo correcto.

Cuando se asume una red de una sola capa oculta, el proceso de entrenamiento parte de ingresar a la red un patrón P de entrenamiento de q componentes. Dicho proceso se puede ilustrar así:

$$P = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_i \\ \vdots \\ P_q \end{bmatrix} \quad (5)$$

La salida final, en función de la entrada neta y de los pesos de la última capa, está dada, entonces, por:

$$a_k^s = f^s \left(\sum_{j=1}^m W_{kj}^s a_j^o + b_k^s \right) \quad (6)$$

A continuación, la salida final de la red se compara con su valor deseado t_k ; de esta manera se determina el error de cada unidad de salida:

$$\delta_k = (t_k - a_k^s) \quad (7)$$

Por tanto, el error de cada patrón que ingresa a la red está dado por:

$$ep^2 = \frac{1}{2} \sum_{k=1}^s (\delta_k)^2 \quad (8)$$

Donde ep^2 es el error medio cuadrático para cada patrón de entrada P ; δ_k es el error en la neurona k de la capa de salida con l neuronas.

El proceso que se acaba de describir debe repetirse para el número total de patrones de entrenamiento (r). Luego se hace necesario minimizar el error medio cuadrático total expresado por:

$$e^2 = \sum_{p=1}^r ep^2 \quad (9)$$

Aquí, e^2 es el error en el proceso de aprendizaje en una iteración, después de haber ingresado los r patrones a la red.

El error determinado r en función de los pesos genera un espacio de dimensiones, donde se refiere al número de pesos de la red. Al tratar de minimizar dicho error en el proceso de aprendizaje, se toma la dirección del gradiente negativo del error en esta superficie, es decir, aquella dirección en la que el error decrece; este es el criterio para la actualización de los pesos:

$$W_{k+1} = W_k - \alpha \nabla ep^2 \quad (10)$$

El gradiente negativo de ep^2 se denota como $-\nabla ep^2$ y se calcula como la derivada del error con respecto a los pesos de conexión de toda la red.

Se puede demostrar que los términos de error para la capa de salida están dados por:

$$\delta_k^s = (t_k - a_k^s) \times f'^s(n_k^s) \quad (11)$$

Este algoritmo es llamado *Backpropagation* o de propagación inversa porque el error se propaga hacia atrás en la red. Por su parte, el error en la capa oculta está dado por:

$$\delta_j^o = f'^o(n_j^o) \times \sum_{k=1}^l \delta_k^s W_{kj}^s \quad (12)$$

Con el valor del gradiente del error se actualizan todos los pesos de la red, empezando por los de la capa de salida:

$$W_{kj}(t+1) = W_{kj}(t) - 2\alpha \delta_k^s \quad (13)$$

$$b_k(t+1) = b_k(t) - 2\alpha \delta_k^s \quad (14)$$

Donde α es la tasa de aprendizaje que varía entre 0 y +1.

Después de actualizar los pesos y ganancias de la capa de salida se hace el mismo proceso con la capa oculta; así se termina el proceso de actualización de pesos, con base en los errores en una red de tres capas:

$$W_{ji}(t+1) = W_{ji}(t) - 2\alpha \delta_j^o P_i \quad (15)$$

$$b_j(t+1) = b_j(t) - 2\alpha \delta_j^o \quad (16)$$

Las expresiones para generalizar este proceso a una red de más capas se pueden deducir fácilmente de la ecuación (12).

2. Series de datos

Ahora bien, para la selección de las variables de entradas, empleadas para el estudio, se tuvieron en cuenta los fundamentales macroeconómicos y las relaciones entre mercados financieros, documentados ampliamente en diversos trabajos como los de Chen, Firth y Rui (2002) o Phylaktis y Ravazzolo (2005). También se consideraron como criterios la disponibilidad de los datos en la periodicidad requerida y el conocimiento de los investigadores sobre el funcionamiento del mercado financiero colombiano.

El conjunto de datos (1493 en total) se presenta en el apéndice 1 y consta de 77 variables², expresadas en términos de precios desde el 21 de junio de 2002 hasta 11 de marzo de 2008. En la gráfica 2 se muestra la serie de precios de la acción de Suraminv para el periodo objeto de estudio.



Fuente: Elaboración propia

² Series históricas obtenidas de Bloomberg.

El paso siguiente en la organización de los datos es la división de la muestra en dos subconjuntos: uno para la estimación del modelo (*in sample*) y otro que no interviene en el cálculo y que se utiliza para la prueba (*out of sample*). Aunque normalmente se acostumbra realizar un ejercicio de partición, para el presente trabajo se efectuaron dos divisiones diferentes, buscando contar con herramientas para realizar un análisis más robusto de los resultados:

- Inicialmente se usó el 68% para estimación *in sample* y se dejó el 32% final de los datos para efectos de pruebas *out of sample*. Esta división corresponde al punto medio aproximado de la caída más significativa de precios de la acción de Suraminv, presentada aproximadamente entre mayo y junio de 2006. Con esto se buscaba probar si utilizando la primera parte de dicha caída para la estimación, los modelos son capaces de predecir *out of sample* la segunda parte.
- En la segunda división se utilizó el 83% de los datos para estimación y el 17% final para pruebas. Con esta división se procuró probar la capacidad predictiva de los modelos luego del cambio estructural que parece haberse presentado en el mercado a partir de la caída grande de precio ya mencionada; sin embargo, se dejó parte de los datos de la nueva dinámica de los precios dentro del conjunto de entrenamiento.

Más adelante se presentan los resultados del modelo, estimados a partir de la división 83% - 17%, por ser el que mejor captura la dinámica de los datos reales fuera de la muestra.

Por último, de acuerdo con las recomendaciones existentes en la literatura sobre RNA, fue necesario, para la estimación de estos modelos *in sample*, subdividir en tres grupos cada una de las muestras: uno de entrenamiento, donde el modelo realizó el proceso de aprendizaje; otro de validación, con el cual se buscó optimizar

el tiempo computacional en el entrenamiento de la red, y un tercero, de prueba, con el que se validó la capacidad de generalización del modelo³.

3. Resultados de predicción

Las redes neuronales exigen la optimización del error de estimación a través de la comprobación de diferentes estructuras, número de capas, número de neuronas por capa, tipo de funciones de transferencia, aprendizaje y entrenamiento. La realización de este proceso implicó el entrenamiento de cientos de redes. Por este motivo, tiene gran importancia la utilización de algoritmos que sean eficientes desde el punto de vista de los tiempos computacionales.

El entrenamiento, a partir de las 77 series descritas, permitió obtener modelos con buen desempeño en términos de predicción. Los mejores resultados se obtuvieron con una estructura de una capa oculta de 11 neuronas, en la cual se empleó la función de transferencia *tansig*, y una neurona en la capa de salida, con la función *purelin*. A la vez, esta red utilizó la función de gradiente descendiente con momento para el aprendizaje, así como el algoritmo *Levenberg-Mardquart* para el entrenamiento.

La tabla 1 resume los estadísticos de evaluación de este modelo, cuyas ecuaciones se pueden observar en el apéndice 2. Se debe resaltar la obtención de un excelente ajuste *in sample* pero, a la vez, de acuerdo a los objetivos inicialmente planteados, se logró un coeficiente R-cuadrado de 0,8358 *out of sample*. Esto refleja que el modelo tiene buena capacidad para capturar la dinámica de los datos reales y ajustarse a ellos en esa parte de la muestra.

³ Esta metodología es ampliamente recomendada en la literatura de RNA. Para una mayor profundización, véase a Yao, Li y Tan (1997).

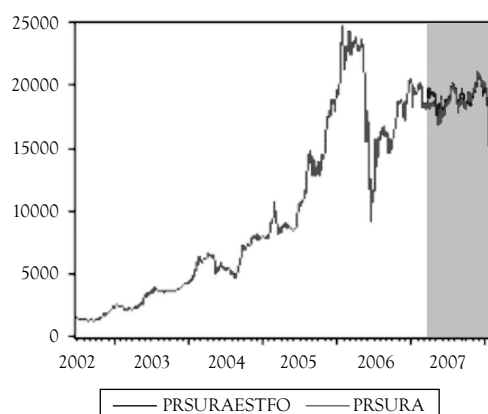
Tabla 1
Estadísticos de error de estimación de la red neuronal

		Redes <i>feedforward</i> 77 precios
In sample	R ²	0,9988
	MSE	53984,8926
	RMSE	232,3465
	RMSPE	0,0324
	MAE	148,9150
	MAPE	0,0217
Out of sample	R ²	0,8358
	MSE	307155,4567
	RMSE	554,2161
	RMSPE	0,0288
	MAE	406,6969
	MAPE	0,0223

Fuente: Elaboración propia

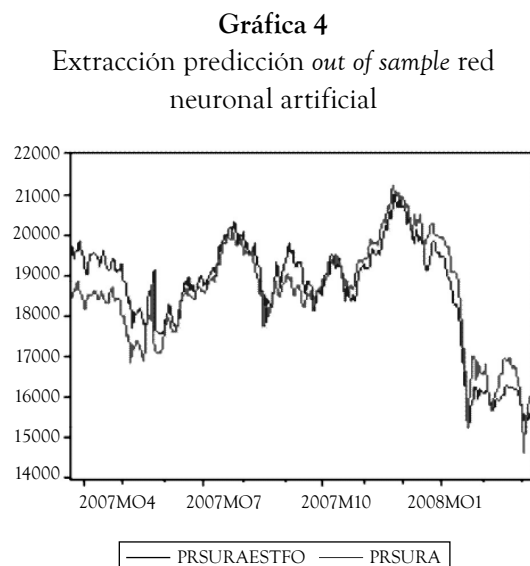
La gráfica 3 deja apreciar, en el área sombreada, la predicción fuera de muestra del modelo y los datos verdaderamente observados. Se resalta que tal como lo muestran los estadísticos de evaluación, se logra capturar gran parte de la dinámica de la serie.

Gráfica 3
Predicción *out of sample*
red neuronal artificial



Fuente: Elaboración propia

La gráfica 4 es una extracción de la gráfica anterior; en ella solo se muestran la predicción y los datos reales para la parte *out of sample* de los datos.



Fuente: Elaboración propia

La evaluación de los residuales del modelo se realizó por medio de la prueba de raíces unitarias *Dickey-Fuller* aumentada, sin intercepto y sin tendencia. Las probabilidades se ubicaron en la región de rechazo de la hipótesis nula de existencia de raíces unitarias. No hay evidencia empírica que muestre falta de estacionariedad en los residuales.

4. Evaluación financiera de las predicciones

Una vez que se tienen las predicciones del comportamiento de la acción de Suraminv en el mercado, por medio del modelo entrenado de redes neuronales, fue importante determinar si por medio de sistemas simples de negociación⁴, elaborados sobre dichas predicciones, hubiera sido posible obtener rendimientos extraordinarios con respecto a una estrategia pasiva, representada en comprar y mantener el *Buy and*

⁴ Un sistema de negociación hace referencia a un conjunto de condiciones programadas para arrojar señales de compra o venta en el mercado.

Hold (B&H), como comúnmente se le conoce. El ejercicio fue realizado con base en los siguientes supuestos:

- Sólo se permiten posiciones *largas*⁵, lo cual es acorde con la regulación del mercado accionario colombiano para el periodo de estudio, en la medida en que legalmente no están permitidas las ventas en corto.
- La estrategia *B&H* y los sistemas alternativos siempre inician, en el primer día del periodo considerado para la comparación, con un valor de portafolio correspondiente al valor de mercado de 100 acciones de Suraminv. Igualmente se considera que todas las compras que realicen los sistemas son de 100 acciones.
- Los faltantes o sobrantes de liquidez del portafolio son financiados a la tasa mínima de expansión del Banco de la República⁶.
- Un costo de transacción que representa la diferencia entre el *Bid* y el *Ask*⁷ del mercado en el momento de realizar las operaciones. Agudelo (2009) estima que este costo para la acción de Suraminv, en el periodo comprendido entre el 24 de abril y el 3 de agosto de 2007, es 0,271%. Si se considera que el 50% de las veces es necesario asumir este costo y las veces restantes lo hace la contraparte, se trabajó con un costo de 0,1355% por operación.

⁵ Nombre que reciben las posiciones de compra en el mercado cuando se está hablando de apuestas direccionales sobre los precios.

⁶ Esta tasa es definida en la página web del Banco de la República (2007) como “la tasa que las entidades financieras deben pagar por tener acceso a la liquidez temporal a través de los repos que se ofrecen por subasta, es decir, con un monto limitado. La tasa a la que se da liquidez a través de la subasta oscila entre esta tasa mínima y la tasa máxima o lombarda de expansión”. Disponible [en línea]: www.banrep.gov.co/informes-economicos/ine_pre_frec3.htm (12/08/2009).

⁷ *Bid* y *Ask* son los nombres que reciben, respectivamente, el mejor precio de demanda y de oferta del mercado en un momento determinado.

- Los modelos se corren cinco minutos cada día antes del cierre del mercado, con la posibilidad, incluso, de realizar operaciones al cierre del mismo día. Para efectos de simplificación, en el ejercicio se supone que los precios existentes cinco minutos antes son *proxies* que responden exactamente a los datos verdaderos de cierre.

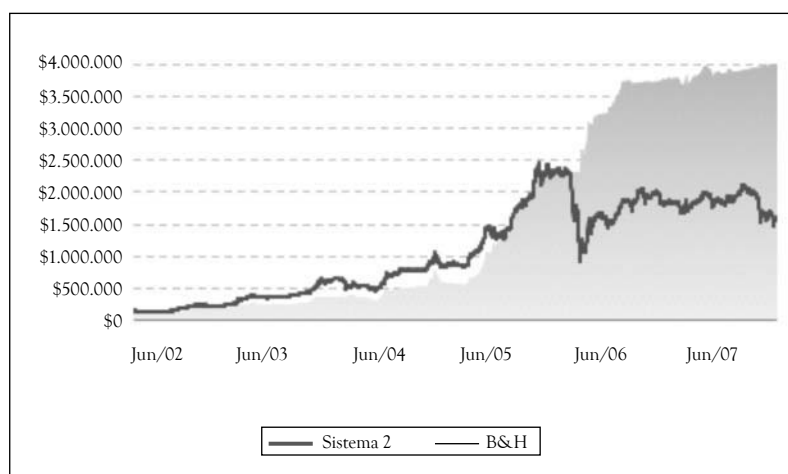
Con base en lo anterior, es necesario resaltar que por la misma lógica que implica la estrategia *B&H*, sus resultados, frente a sistemas de negociación que puedan entrar y salir del mercado, se ven fuertemente favorecidos en épocas de tendencia alcista, mientras que se pueden ver perjudicados altamente en movimientos laterales o bajistas. Por esto, un buen sistema de negociación para mercados con fuertes tendencias debería seguir muy de cerca la estrategia *B&H* durante épocas alcistas de precios y ganarle de forma importante en movimientos laterales o bajistas.

La estrategia de negociación, empleada y denominada como *Sistema*, trabaja bajo las siguientes condiciones:

- No hay posición. Se evalúa si ante un cumplimiento exacto de la predicción para la siguiente jornada, la serie de precios de cierre cruza hacia arriba el promedio móvil ponderado⁸ de tres días de dicha serie. En caso de presentarse esta situación, se genera una señal de compra al precio de cierre del día.
- Hay posición abierta. El cruce se presenta en sentido contrario y se genera, entonces, una señal de venta.

El número de días del promedio móvil ponderado es el resultado de un proceso de optimización en el que se buscaba maximizar los rendimientos del *Sistema* en la parte *in sample* de los datos. La gráfica 5 presenta la evolución del valor de los portafolios para este ejercicio.

Gráfica 5
Valoración portafolio *Sistema* y *B&H*



Fuente: Elaboración propia

⁸ Los promedios móviles ponderados son usados ampliamente como herramientas de análisis técnico de mercados. Se calculan mediante una ponderación lineal de los precios, multiplicando el primero (más reciente) por n , el segundo por $(n-1)$, el tercero por $(n-2)$ y así sucesivamente, donde n es la periodicidad del promedio móvil. El resultado final es dividido por la suma de las ponderaciones.

El comportamiento observado es precisamente el que se espera de un buen sistema de *trading* en épocas de tendencia al alza, ya que solo perdió partes relativamente pequeñas del movimiento. Incluso, en 2005, año de fuerte tendencia positiva, el *Sistema* logró superar a la estrategia *B&H*. En términos agregados para todo el periodo de estudio, el sistema de negociación alcanzó un rendimiento anualizado de 78,27% frente al 51,23% que obtuvo la estrategia *B&H*.

En el periodo comprendido entre principios de 2006 y marzo de 2008, caracterizado por movimientos a la baja y laterales, se observa que el *Sistema* consiguió una ventaja muy importante frente a la estrategia *B&H*, consistente durante todo este tiempo.

De igual manera, se realizó el análisis comparativo para la parte *out of sample* de la serie. Los resultados mostraron que nuevamente el *Sistema* es muy superior a la estrategia *B&H*, en la medida en que es capaz de mantener el valor del portafolio prácticamente constante en una época de movimientos laterales y bajistas del mercado.

A esta altura, es conveniente detallar la evolución del valor de portafolio en diferentes subperiodos que permitan evaluar si la ventaja que presenta el *Sistema* frente a la estrategia *B&H* se mantiene de forma consistente. En tal sentido, la tabla 2 presenta un resumen de los rendimientos anualizados para todo el periodo de estudio, detallando para cada uno de los años que lo componen y para sendas partes, *in sample* y *out of sample*, de los datos.

Tabla 2
Rendimientos anualizados
Sistema y estrategia B&H

Año	Sistema	B&H
2002	65,98%	118,35%
2003	27,85%	87,14%
2004	59,05%	86,17%
2005	176,68%	146,35%
2006	94,05%	4,17%
2007	3,77%	-2,56%
2008	50,99%	-65,23%
<i>In Sample</i>	96,82%	69,20%
<i>Out of Sample</i>	2,80%	-14,20%
jun. 2002 – mar. 2008	78,27%	51,23%

Fuente: Elaboración propia

En términos agregados, se observa que el desempeño promedio del sistema de negociación, si se tienen en cuenta las diferentes tendencias del mercado, es superior al de la estrategia *B&H*: en las épocas de subida tiene un comportamiento ligeramente inferior al de la estrategia pasiva, pero en épocas de movimientos laterales o bajista obtiene rendimientos muy superiores. En la medida en que se involucren mayor historia de los precios de la acción y una mayor cantidad de movimientos alcistas, laterales y bajistas, se espera que la ventaja del *Sistema* sea aun más marcada, ya que la muestra del presente estudio estuvo caracterizada por una fuerte tendencia al alza en la mayor parte del tiempo.

Conclusiones

La investigación permitió contrastar la teoría de eficiencia débil de mercado en la acción de Suramericana de Inversiones S.A. (Suraminv), acción líder en el mercado bursátil colombiano durante el periodo de estudio: junio de 2002 a marzo de 2008. Los resultados arrojaron como conclusión que en este mercado, contrario a la

hipótesis defendida por la teoría mencionada, sí es posible hacer predicciones confiables y obtener rendimientos extraordinarios con respecto a una estrategia *Buy & Hold*. El trabajo fue más allá que muchos otros de su estilo, en el sentido de que además de lograr un buen pronóstico *in sample*, demostró que es posible obtener buenas estimaciones del comportamiento de la acción *out of sample*; por tanto, evalúa de manera más confiable la verdadera posibilidad de aprovechamiento financiero de los resultados.

A partir de 77 variables expresadas en términos de precios y tratando de controlar la posibilidad de *data snooping*, se ejecutó una división de la muestra en dos grupos: el primero, *in sample*, se usó para el ajuste de los modelos, y el segundo, *out of sample*, que no interviene en el entrenamiento, se empleó para evaluar la capacidad predictiva de los modelos. Después se valoraron diversas estructuras de redes *feedforward*, con base en el algoritmo de entrenamiento *backpropagation*.

Los resultados de las redes mostraron que es posible capturar adecuadamente la dinámica de los datos *in sample* y *out of sample*. Tal como se planteó desde el principio, el trabajo no se quedó solamente en un ejercicio de predicción, también regresó al campo financiero en la medida en que teniendo en cuenta costos de transacción y oportunidad, evaluó la posibilidad de aprovechar esa predictibilidad desde el punto de rendimientos de exceso, con miras a contrastar la teoría de eficiencia débil de mercado. Dicha valoración se realizó por medio de la comparación, en diferentes periodos, entre sistemas de negociación basados en las predicciones de la red de mejor resultado y la estrategia pasiva de comprar y mantener.

En la evaluación financiera de la predicción resultó que, contrario a lo que defiende la teoría de eficiencia débil de mercado, es posible obtener, por medio de los sistemas de nego-

ciación en el largo plazo, rendimientos muy superiores a los de la estrategia pasiva con un nivel alto de confianza. En tal sentido, es importante hacer énfasis en que, en ciertos subperiodos de alta tendencia alcista de los precios, la estrategia pasiva produce rendimientos superiores. Sin embargo, al calcular a nivel agregado las diferentes tendencias que se presentaron durante el periodo de estudio, el desempeño promedio del sistema de negociación se evidenció como superior.

Bibliografía

- Acosta, M.; Salazar, H. y Zuluaga, C. (2000). "Tutorial de redes neuronales". Universidad Tecnológica de Pereira, Facultad de Ingeniería Eléctrica, <http://ohm.utp.edu.co/neuronales/> (12/08/2009).
- Agudelo, D. (2009). "Costos de transacción y medidas de liquidez intradiaria en BVC". Trabajo presentado en el VI Simposio Nacional y III Internacional de Docentes en Finanzas, Bogotá, Colombia.
- Agudelo, D. & J. Uribe. (2007). *¿Realidad o sofisma? Poniendo a prueba el análisis técnico en las acciones colombianas*. Medellín: Universidad EAFIT.
- Álvarez Díaz, M. & Alvarez, A. (2003). "Predicción no-lineal de tipos de cambio: algoritmos genéticos, redes neuronales y fusión de datos." Disponible [en línea]: <http://economiaaplicadavigo.org/dt/2003/dt0301.pdf> (12/08/2009).
- Aristizábal, M. (2006). "Evaluación asimétrica de una red neuronal artificial: aplicación al caso de inflación en Colombia", *Borradores de economía*, 377. Banco de la República de Colombia, pp. 1-52.
- Chen, G.; Firth, M. & Rui, O. (2002). "Stock market linkages: evidence from Latin America", *Journal of Banking & Finance*, 26. United Kingdom: Elsevier, pp. 1113-1141.
- Chen, A.; Leung, M. & H. Daouk. (2003). "Application of neural networks to an emerging financial market: forecasting and trading the

- Taiwan stock index", Computers and Operations Research, 30, pp. 901-923.
- Concepto tasa mínima de expansión, [http://www.banrep.gov.co/informes-economicos/ine_pre_frec3.htm\(12/08/2009\)](http://www.banrep.gov.co/informes-economicos/ine_pre_frec3.htm(12/08/2009)).
- Fama, E. F. (1970). "Efficient capital markets: a review of theory and empirical work", *Journal of Finance*, 2(25). American Finance Association, pp. 383-417.
- Fernández Rodríguez, F., González Martel, Ch. y Sosvilla Rivero, S. (2000). "On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market". *Economics Letters* Vol. 69, pp. 89-94.
- Franses, P.H. & van Dijk, D. (2000). *Non-Linear Time Series Models in Empirical Finance*. Primera Edición. United Kingdom: Cambridge university press.
- Gil Zapata, M & Pérez Ramírez F. (2005). Análisis y predicción de la acción de la empresa acerías paz del río utilizando un modelo garch(1,1) y redes neuronales artificiales. *Revista de Ingenierías Universidad de Medellín*, Vol. 4, No. 007. Universidad de Medellín. Medellín, pp. 83-97.
- Haykin, S. (1999). *Neural networks: a comprehensive foundation*. Segunda Edición. New Jersey, USA: Prentice Hall.
- H. Yann & Y. Hua, "On rank of block Hankel matrix for 2-D frequency detection and estimation". (1996). *IEEE Transactions on Signal Processing*, Vol. 44, No. 4, pp. 1046-1048.
- Jalil Munir, A. & Misas, M. (2007). "Evaluación de pronósticos del tipo de cambio utilizando redes neuronales y funciones de pérdida asimétricas" *Revista Colombiana de Estadística*, Vol.30 No. 1, pp. 143-161.
- Matlab. (1998). *Neural network toolbox User's Guide*. U.S. : The MathWorks.
- Mcnelis, P. D. (2005). *Neural networks in finance: gaining predictive edge in the market*. United Kingdom: Elsevier.
- Pantoja Rojas, M. (2000). "Comparative analysis of time series forecasting with neuronal networks, arima models and garch process for non-stationary time series". *rev.ing*. Disponible [en línea]. 2004, No. 20, pp. 108-113.
- Parisi F, Antonino & Parisi F, Franco & Guerrero C., José Luis. (2003). "Modelos predictivos de redes neuronales en índices bursátiles," *El Trimestre Económico*, Fondo de Cultura Económica, vol. 0(280), pp. 721-744.
- Phylaktis, K. & F. Ravazzolo. (2005). "Stock market linkages in emerging markets: implications for international portfolio diversification", *Journal of International Financial Markets*, 2(15). Institutions and Money, pp. 91-106.
- Roberts, H. (1967). "Statistical versus clinical predictions of the stock markets". Unpublished manuscript, Center for research in Security Prices, University of Chicago.
- Sallehuddin et al. (2007). "Forecasting time series data using hybrid grey relational artificial neural network and auto regressive integrated moving average model." *Neural Network World* 6, 07, pp. 573-605.
- Shachmurove, Y. (2000). "Utilizing artificial neural network model to predict stock markets". Working Paper # 00-11, Caress.
- Villada et al. (2008). "Pronóstico del precio de la energía eléctrica usando redes neuronales artificiales" *Rev. Fac. Ing. Univ. Antioquia*. No. 44, pp. 111-118.
- Wong, B. K. (1995). "A bibliography of neural network business application research: 1988-September 1994", *Expert Systems*, 3 (12). United Kingdom, pp. 253-261.
- Yao J.; Li, Y. & Tan, C. (1997). "Forecasting the exchange rates of CHF vs. USD using neural networks", *Journal of Computational Intelligence in Finance*, 2(5), pp. 7-13.

Apéndice 1
Series de datos de entrada para los modelos

Acciones Colombianas	Suramericana	Monedas	TRM
	Acerías Paz del Río		Euro
	Banco de Bogotá		Yen
	Bancolombia O		Libra
	Bancolombia P		Franco suizo
	ADR – Bancolombia		Real
	Colombiana de Inversiones		Peso mexicano
	Coltejer		Peso chileno
	Chocolates		Peso argentino
	Éxito		Won surcoreano
	Fabricato		Dólar Taiwán
	Grupo Aval		Rublo ruso
	ISA		Yuan chino
	Tablemac		Dólar canadiense
	Valorem		Dólar australiano
	Cemargos		Lira turca
	Inverargos		
Índices Accionarios	IGBC (Colombia)	Tasas	Tasa de la FED
	Dow Jones (EUA)		Tasa BanRep
	S&P 500 (EUA)		Tasa Banco de Japón
	NASDAQ COMPOSITE (EUA)		Tasa del Banco Central Europeo
	S&P EUROPE 350 (Europa)		Tasa del Banco del Reino Unido
	FTSE 100 (Reino Unido)		Tasa de Brasil
	CAC 40 (Francia)		Tasa genérica tesoros EUA 10 años
	DAX (Alemania)		LIBOR 1 Semana
	NIKKEI 225 (Japón)		LIBOR 1 Mes
	SHANGHAI COMPOSITE (China)		LIBOR 3 Meses
	HANG SENG (China)		LIBOR 6 Meses
	SENSEX 30 (India)		ITES
	BOVESPA (Brasil)		DTF
	BMV (México)	Commodities	Crudo de petróleo WTI
	MERVAL (Argentina)		Crudo de petróleo Brent
	IPSA (Chile)		Precio del café colombiano
	IGBVL (Perú)		Oro
Futuros	Futuro Genérico Dow Jones		E-mini crudo de petróleo
	Futuro Genérico S&P		Futuros del precio del café
	Futuro Genérico Nasdaq	Otros	VIX
			EMBI Global y EMBI+ Colombia
			Spread deuda colombiana
			UVR

Apéndice 2

Estadísticos de error para evaluación de los modelos

- R-Cuadrado:

$$R^2 = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y}_t)^2}$$

- Estadístico del error medio cuadrático (MSE):

$$MSE = \frac{1}{2} \sum_{t=1}^n (\hat{y}_t - y_t)^2$$

- Estadístico de raíz del error medio cuadrático (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2}$$

- Estadístico de raíz del error porcentual medio cuadrático (RMSPE):

$$RMSPE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left(\frac{\hat{y}_t - y_t}{y_t} \right)^2}$$

- Estadístico del error medio absoluto (MAE):

$$MAE = \frac{1}{n} \sum_{t=1}^n |\hat{y}_t - y_t|$$

- Estadístico del error porcentual medio absoluto (MAPE):

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{y_t} \right|$$